DOCUMENT RESUME

ED 058 313 TM 001 028

AUTHOR	Reilly, Richard R.
TITLE	Critical Incidents of Graduate Student Performance.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	TM-1
PUB DATE	Apr 71
NOTE	14p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Academic Performance; *Critical Incidents Method; *Graduate Students; *Graduate Study; *Student Behavior; *Student Evaluation

ABSTRACT

A sample of faculty members from three fields of study (English, Psychology, Chemistry) responded to a survey eliciting descriptions of specific observable incidents which caused them to change their estimate of the competence of a graduate student. A listing of the incidents retained from the responses, as well as copies of the initial letter of inquiry, response form, and follow-up letter are included. (DLG)

ED058313

GRADUATE RECORD EXAMINATIONS BOARD
Technical Memorandum No. 1

CRITICAL INCIDENTS OF GRADUATE STUDENT PERFORMANCE

by

Richard R. Reilly

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Educational Testing Service
Princeton, New Jersey

April 1971

TM 001 028

CRITICAL INCIDENTS OF GRADUATE STUDENT PERFORMANCE

Researchers seeking to demonstrate the validity of test scores or other information for predicting graduate school performance have encountered three major difficulties. First, the small sample sizes available at the graduate level make results, especially when several predictors are involved, subject to a considerable degree of error. Second, the fact that students within a given department have gone through an elaborate screening process, and as a result are usually quite homogeneous with respect to most predictor information, often leads to restricted variation in the predictor score distributions. Finally, there is the difficulty of establishing an adequate criterion of graduate school performance. Grade point average (GPA), while it has been the most widely used criterion, has also been the most severely criticized. Perhaps the most important and valid of these criticisms is that the GPA represents only a limited aspect of graduate school performance (Lannholm et al, 1968). The alternative criterion of attainment versus nonattainment of the doctorate is logically appealing, but a substantial period of time must elapse before such data can be collected. Another problem with this pass/fail criterion is that it lacks sensitivity, since it cannot take into account the various qualitative levels of performance of individuals attaining (or not attaining) the Ph.D. In an effort to procure more sensitive data, some investigators have used global ratings of success, generally employing a single a priori rating scale. Most such scales, because they are not related to observable events, are subject to the common sources of rater error (e.g., leniency, central tendency, etc.), and ignore the possibility that several distinct dimensions of graduate performance might exist.

A study by Hilton, Kendall, and Sprecher (1970) describes an attempt to develop an extensive set of rating scales for use in the graduate business school setting. First, 15 qualities necessary for success in graduate business study were posited. Business school faculty members were then asked to provide specific examples of behavior illustrating each quality. Finally, a behaviorally-anchored rating scale was developed for each of the 15 qualities. Although the study was limited in that it was restricted to graduate business schools, it is notable in that the effort was made to relate judged performance to specific and observable events.

Any method of obtaining judgmental performance ratings is subject to error, of course. The more common errors such as the halo effect or excessive leniency may result, in part, from speculations and inferences on the part of the rater as to what sets of behaviors or events are encompassed by each trait or quality being rated.

While speculation and inference can probably never be entirely eliminated where judgments of complex human performance are being made, it should be possible to reduce rater errors to a considerable extent by requiring raters to deal only with specific and observable incidents or events. The critical incident technique (Flanagan and Burns, 1955) is based on this guiding principle and has been used as a tool in evaluating

the performance of diverse groups, ranging from salesmen (Kirchner and Dunnette, 1957) to factory workers (Flanagan and Miller, 1955).

The present study was an attempt to collect descriptions of specific observable events which caused graduate educators in each of 3 fields of study to raise or lower their estimate of the competence of graduate students. It was hoped that the results would constitute an initial step toward defining several dimensions of graduate student performance.

METHOD

A total of 150 faculty members -- 50 from each of the fields of English, Psychology and Chemistry -- comprised the sample. Names of professors were chosen from listings appearing in The Annual Guides to Graduate Study (Hegener, 1970). Although some attempt was made to sample a representative range of institutions, no strict randomization procedures were followed. The 3 fields chosen were felt to be fairly representative of the range of graduate disciplines.

Each educator in the sample was sent a letter describing the study, and asking him to provide specific incidents which, when they occurred, caused him to either lower or raise his estimate of the competence of a graduate student. A follow-up letter was sent to all nonrespondents one month after the original request. Copies of these letters, and of the response form, are included in the Appendix.

Returns were examined and tabulated, and a final list of incidents was prepared. Three criteria were used to decide whether an incident would be retained for the final list: (1) The incident should be neither too general nor too specific; (2) it should be applicable across disciplines or fields; (3) it should not be highly similar to another incident on the list. Since response styles varied considerably, incidents were edited where necessary for the sake of consistency.

RESULTS AND DISCUSSION

A total of 99 faculty members -- almost two-thirds of those sampled -- responded to the survey. (Only 75 of the response forms returned were judged usable, however, with the bulk of the nonusable returns consisting of misinterpretations and failure to comply with the questionnaire directions.) A total of 336 separate incidents were provided. The number of returns, and the mean and modal number of incidents for each discipline area, are given in Table 1. The modal number of 6 incidents probably resulted from the format of the response form, which provided space for 6 (see Appendix). The 52 incidents finally retained are shown in Table 2. The order of the items was randomly determined, since the same list will be used as a rating device in the future phases of this research. For the same reason, an attempt was made to balance the numbers of positively and negatively stated incidents.

It is tempting to try to derive logical factors or dimensions of graduate student performance by simply grouping items that seem to reflect the same underlying quality. "Initiative," as one such factor, could be defined by items 2, 18, 20, 30, 32, 36, and 51. Except for this example, however, the temptation was resisted, since it was felt that the most meaningful and useful factors can be derived empirically.

Two characteristics of the list are worth noting First, many of the items either explicitly or implicitly involve research methodology. Educators in all three of the disciplines sampled appeared to share a concern with research -- which is reasonable since the Ph.D. is by definition a research degree. Second, a few items are specific to fields where experimentation or laboratory work is involved (e.g., item 24). These items were included because of their apparent importance, as suggested by the frequency with which they were cited by both Psychology and Chemistry faculty members, and because the majority of doctorates awarded are in fields such as the physical and biological sciences (Cartter, 1965) where such items are applicable.

Because the sample was not a strictly random one no attempt was made to draw generalizations about differences in critical behaviors among the three separate discipline areas studied. It was felt that the main purpose of this study was to gather a number of critical incidents to serve as a basis for further study, and that this was achieved.

CONCLUSIONS AND RECOMMENDATIONS

The results of this study may be viewed as a successful initial step toward empirical definition of several criteria of graduate student performance. Except for a few items dealing with experimentation or laboratory work, the incidents obtained appear to be generally applicable across disciplines.

The present report can be regarded as one limited phase of a more broadly conceived research plan aimed at better definition and greater understanding of criteria of graduate student performance. While such a plan must of necessity be rather flexible, several potential studies based on, or related to, the results of the present study can be mentioned.

One immediately useful research project would involve having a representative sample of graduate educators from the fields sampled in the present study use the list of critical incidents developed in this study to rate (each of their) graduate students. Each incident would be accompanied by a three-point scale allowing the rater to indicate whether he had observed the incident, had observed its contrary, or had had no opportunity to observe it. Given adequate sample size, factor-analytic methods could be employed to empirically derive criterion dimensions. The examination of simple item statistics (e.g., means and variances) across and within discipline areas, might result in other useful outcomes.

Although an attempt was made in the present study to choose disciplines representative of the range of areas in which graduate programs are offered, there is certainly no assurance that the list covers the important aspects of graduate student performance in all fields of study. Thus, another study might extend the present methods to a wider sampling of disciplines, to determine whether the present list should be modified or augmented in order to be generally applicable.

Once a reasonably stable set of criterion dimensions has been identified, focus can be shifted to determining how to best predict performance on each factor. Another study, or series of studies, might examine the usefulness of the GRE, as well as other information, for predicting success along each of the several criterion dimensions.

TABLE 1

Summary of Returns by Discipline

Discipline	Total Number of Usable Returns	Mean Number of Incidents per Return	Modal Number of Incidents per Return
English	21	4.5	6
Psychology	29	4.6	6
Chemistry	25	4.5	6
Total	75	4.5	6

TABLE 2

1. Rigidly followed a research plan when more flexibility would have been advantageous.

2. This student's willingness to pursue unassigned readings was reflected by a broader than average knowledge of most topics.

3. Consistently offered well founded and constructive criticisms of other students' presentations.

4. When making a judgment or reaching conclusions this student supported his position with carefully documented research.

5. Was able to master a difficult research technique in an unusually short period of time.

6. During informal discussions with faculty this student displayed a genuine interest in and commitment to his field.

7. Was careless in reporting data.

8. In conducting research this student relied too heavily on one particular research tool.

9. Displayed an inability to write competently.

10. Failed on one or more occasions to complete a major assignment on time.

11. In writing a report this student synthesized material from two independent fields.

12. Showed himself to be unfamiliar with a major research tool in his field.

13. Was able to consider several markedly different approaches to a research problem and view them objectively before choosing one.

14. Performed an experiment without making proper checks.

15. Handles even the most menial assignment (e.g., paper grading) with care and responsibility.

16. Talks at great length in class but exhibits little understanding of material on papers and tests.

17. Was unable to formulate a testable hypothesis from a theoretical analysis.

18. Developed an original way of handling a research problem.

19. When presenting a paper handled a difficult topic with considable skill.

20. Became more proficient in a useful outside field under his own initiative.

21. Became distracted by nonacademic, nonprofessional interests.

22. Was heavily dependent on direction from faculty and appeared unable to undertake any independent investigation.

23. Devoted considerable time to helping other students with problems.

24. Exhibited carelessness with laboratory equipment.

25. Although able to criticize studies with facility was unable to suggest better alternatives.

26. Despite discouraging advice from faculty this student pursued his interest or ideas and was successful.

27. Seldom, if ever, engages in informal contacts with faculty or fellow graduate students.

28. Was often unable to consider new ideas objectively because of strongly held prejudices.

29. Repeatedly made irrelevant remarks during class or seminar discussion.

30. This student usually did more than the required work.

31. Submitted a report which was incomplete.

32. Independently planned and executed a study which made a worthwhile contribution to his field.

33. Was unable to effectively apply a research technique.

34. This student displayed a familiarity with the latest developments in his field.

35. Showed imagination and originality in teaching a traditionally dull topic to an undergraduate class.

36. Learned an important research skill on his own.

37. When this student asked a question it was always relevant and usually perceptive.

38. Became quickly and enthusiastically involved in a research project.

39. Was unwilling or unable to accept criticism.

40. Did not hesitate to repeatedly ask questions of faculty until he fully understood an issue.

41. Submitted a paper or report which failed to address the assigned issues.

42. Stimulated great interest and enthusiasm in undergraduate courses in which he was an instructor.

43. Was able to articulately defend his position and ideas.

44. Avoided challenging courses or work.

45. Conducted a data analysis which was inappropriate for the experiment as designed.

46. Attempted to carry out poorly planned research.

47. Despite one or more setbacks continued to work on research until it was successfully completed.

48. Was unprepared for a seminar.

49. Submitted a paper which merely summarized what he had read.

50. Showed an ability to examine carefully an author's premises and frame of reference before accepting conclusions.

51. Asked for more work when none was assigned.

52. Presented ideas in a seminar, paper, or test in a poorly organized and disjointed fashion.

References

Cartter, A. M. An Assessment of Quality in Graduate Education Washington, D.C. American Council on Education, 1966.

Flanagan, J. C. and Burns, R. K. The employee performance record: a new appraisal and development tool. Harvard Business Review, 1955, 33(5), 95-102.

Flanagan, J. C. and Miller, R. B. The Performance Record: Handbook for Foremen and Supervisors of Hourly Employees. Chicago, Illinois: Science Research Associates, 1955.

Hegener, Karen (Ed.) The Annual Guides to Graduate Study. Peterson's Guides, Inc., Princeton, New Jersey.

Hilton, T. L., Kendall, L. M. and Sprecher, T. B. Performance criteria in graduate business study: parts I and II -- development of rating scales, background data form, and the pilot study. Research Bulletin 70-3. Princeton, New Jersey: Educational Testing Service, 1970.

Kirchner, W. K. and Dunnette, M. D. Identifying the critical factors in successful salesmanship. Personnel, 1957, 34(2), 54-59.

Lannholm, G. V., Marco, G. L. and Schrader, W. B. Cooperative studies of predicting graduate school success. Graduate Record Examinations Special Report 68-3. Educational Testing Service, Princeton, New Jersey.

APPENDIX

EDUCATIONAL TESTING SERVICE PRINCETON, N.J. 08540

Area Code 609
921-9000
CABLE-EDUCTESTSVC

Developmental Research Division

As part of its continuing interest in research in graduate education, the Graduate Record Examinations Board is sponsoring a small research project to explore more systematic and better ways of evaluating graduate student performance. As the first stage in this project we are asking for the cooperation of a selected number of individuals who have taught or are teaching at the graduate level. You are asked to participate by taking a few moments to fill out and return the attached form (a self-addressed envelope is provided).

In this survey we are asking established graduate educators to provide us with their insights into what behavior distinguishes the outstanding or highly successful graduate student from the less successful one. One way of doing this is to recall specific incidents which, in the past, have caused you to make a shift in your judgment about the relative effectiveness of a graduate student. That is, recall occasions which have caused you to either increase or decrease your estimation of the competence of a student.

The following two examples may help to clarify what is meant:

(1) "Displayed an unusual tenacity in attacking a research problem."

(2) "On repeated occasions asked pointless questions during seminar discussions."

We would like you to provide us with similar incidents from your own experience. Please fill out the attached form (list as many incidents as you wish) and return by October 30, 1970. Since we are surveying a limited number of educators in this initial phase your cooperation is extremely important and will be highly appreciated.

Very truly yours,

Richard R. Reilly
Associate Research Psychologist

Name_____Department_____

In the spaces provided below briefly describe several specific incidents (as many as you wish), which, in the past, have caused you to raise or lower your estimation of the competence of a graduate student. Please return the completed form in the self-addressed envelope.

1.

2.

3.

4.

5.

6.

(Use other side if necessary)

EDUCATIONAL TESTING SERVICE PRINCETON, N. J. 08540

Area Code 609
921-9000
CABLE-EDUCTESTSVC

Developmental Research Division

November 6, 1970

Dear Professor

Recently, a letter was sent regarding your participation in a survey of graduate educators. If you did not receive, or have lost this letter, I have attached a copy along with the survey form.

I hope you can find a few moments to fill out and return the form. If you cannot, please at least sign and return the blank form in the self-addressed envelope so that we know you received this letter.

Very truly yours,

Richard R. Reilly
Associate Research Psychologist

RRR/jh
Attachments